A Numbers Game: How Analytics are Affecting Decision-Making in Professional Baseball

A Senior Honors Thesis

Submitted in Partial Fulfillment of the Requirements
for Graduation in the Honors College

By
Paul Cifonelli
Journalism, Broadcasting and Public Relations Major

The College at Brockport
December 11, 2020

Thesis Director: Kimberly Young, Lecturer, Journalism, Broadcasting and Public Relations

*Educational use of this paper is permitted for the purpose of providing future
students a model example of an Honors senior thesis project.*

Table of Contents

Abstract

The game of baseball has been based on numbers since the first professional game was played. The game has evolved as the use of statistics has evolved and vice versa. At this point in the baseball's progression, analytics and numbers are affecting every decision that is made, whether it is on the field or in the front office.

This is not a concept that can be depicted using numbers. To better get at this concept, and to use skills taught at SUNY Brockport, journalism would be the best way to fully understand how teams use analytics and what impact those numbers have on the game.

Journalism allows for the human element of things to be more pronounced. It is a form of writing that puts the subject first and takes the writer out of play. Due to this, having professionals to give their input on the topic allows experts their opportunity to say exactly how things happen. It takes guess work out of play and lets the story be told by someone who lives it out on a day-to-day basis.

Acknowledgements

First and foremost, I need to thank my thesis advisor, Kim Young. Without your help, guidance and encouragement, none of this would have been possible. The excitement you have shown from the very beginning of this project kept me motivated, even when things looked bleak.

The next two people I need to thank are Sanjay Choudhury and Alex Nakahara. When everything felt lost for me, they were the two people who agreed to give a little time and insight into their jobs. While it was not a lot, it meant a lot and helped this project get to the point it is at.

I also want to thank my family for their support. They listened when I was frustrated and did not know what to do and provided me ideas in those times, even if I did not want to hear it at the time.

Additionally, I need to thank my friends. They are the ones who share my passion for baseball with me most often and were just as excited to hear about this project as I was to complete it.

Lastly, I want to thank all the professors I have had in the Journalism, Broadcasting and Public Relations department. You have put in a lot of time and effort to helping me to be where I am today, and I am very grateful for that. One of the goals I had for this project was to make the department proud as a thank you for all the effort everyone has put forward to make sure I succeed.

Thesis Body

Methods

This project has been complex and hit many roadblocks, so breaking it down into a methods section and a reflection section will make it easiest to explain and understand. The idea for this project came about during my senior year in high school when I was deciding what I wanted to major in. Initially, I was a double major in journalism, broadcasting and public relations and mathematics. I wanted to be able to use a knowledge of statistics and models in my writing, primarily centered around baseball. A natural interest in how decisions were made and how numbers impact the game developed, then the idea for using this as my thesis topic became clear.

The initial thesis project that was proposed in HON395 was a multimedia story with elements of writing, videography and photography. A multitude of things messed with this idea and made it unattainable, primarily the COVID-19 pandemic.

With the original idea, I reached out to Jason Smorol, the general manager of the Syracuse Mets, a minor league baseball team. In January of 2020, Smorol agreed to be interviewed for the story and allow me access to the field and clubhouse before a game. He would make sure I could get interviews with a few players and coaches and let me take as many photos and videos as I would need for the story. From there, my plan would be to have the people I interviewed connect me with others, hopefully players who were actively playing in the major leagues. The major league players have the most access to advanced data of all players, so they would have been the ideal people to talk to about how analytics impact the game on a day-to-day basis. Additionally, I was hoping that Smorol would have been able to connect me with the New York Mets' front office, as the Syracuse Mets are owned by the New York Mets.

I also reached out to Bill James, one of the godfathers of modern-day baseball analytics. He developed statistics that are now commonly used to evaluate players and worked in the Boston Red Sox organization for many years. Through his website, Bill James Online, I received a response that James was booked for the next few months but that I would be placed on a waiting list to speak with him. Unfortunately, I never heard back from James or any of his representatives, despite another email asking for a meeting.

The original plan came to a screeching halt when the COVID-19 pandemic began. Things became chaotic in everyone's world. Students were learning how to learn online, athletes were looking for new places to work out and keep in top shape and executives needed to figure out if a season was going to be played. I did not hear from Smorol again until early September. At this point, the minor league baseball season was cancelled and the major league season had just gotten underway, with an abbreviated, 60-game season taking place. It was then that he told me he alone would not be able to help because he did not work closely with any analytics. He also did not connect me with the New York Mets, so I needed a new plan.

I decided to turn my attention to the players. While executives are the ones who use analytics to make decisions, everything in baseball ultimately comes down to the players on the field. The story could turn from an overarching one covering all members of the baseball community to just the players. This proved to be a more difficult task than the first one.

Players' contact information is not readily available, as people would be contacting them all the time to give their two cents on their recent performances or actions off the field. The one place players and the public all coexist is social media, so I made an attempt to get in contact with some players there.

There were two problems with this plan. The first problem was that all players who have played at the major league level are verified. This includes retired players. When someone is verified on Instagram or Twitter, they will not see direct messages unless they allow a person to send one to them. In order to try and get their attention and get them to see my direct message, I was waiting to send the message until they made a post. Once they posted, I sent the direct message and commented on the post. The comment was a condensed version of the request I made in the direct message, explaining who I was and what I wanted. In multiple cases, I was able to see that the player viewed my direct message, but never responded to it. After about a month of fruitlessly trying this, I knew I needed to make one last ditch effort.

At this point, it was the beginning of November and I was running out of ideas. I decided to send emails to all 30 major league teams and see if I could get any responses. It felt like a long shot, but it was the last step that I felt I could take to save the project.

I received responses from two teams, the Toronto Blue Jays and the Philadelphia Phillies. I worked with a media manager over email from the Blue Jays. After a few emails, I was connected with Sanjay Choudhury, the team's assistant director of research and development. For the Phillies, whoever is in charge of their media relations sent my email directly to Alex Nakahara, the organization's senior quantitative analyst.

Once I was in contact with Choudhury and Nakahara, I was able to secure interviews with them at the end of the week. Prior to the interviews, I needed to do my research on the organizations, the interviewees and their positions. After conducting my research, I developed some questions specific to each interviewee and some that were general and could be asked to either Choudhury or Nakahara.

The interviews were held over Zoom and took about a half hour each. I am typically an off-script interviewer, meaning that I usually do not have any preprepared questions. I know where I want to take the story before I begin an interview and know how the interviewee can help with the story. I compare writing a story to a road. There will be flat land where you, the writer, can tell the story. Sometimes you will need help and assistance from other people, which I equate to going over a bridge or through a tunnel. The interviewees are there to fill in the holes or build a bridge between two ideas or facts.

With this story, I wanted to be more sure than ever that I had all the information I needed. Considering my success rate at getting responses from the teams, I needed to have a plan, which I was not used to. The interviews still went smooth, but I found myself spending less time in full conversation with Choudhury and Nakahara because I was so focused on getting as much information out of them as possible in as short a time as I could. I was afraid they would have to leave for a meeting or some other obligation mid-interview, so I wanted to ask all my questions as early in the interview as possible.

Despite it being a different style of interview than I'm used to, I still think they went well. I had some great discussions with them and, due to how closely I follow baseball, was very knowledgeable on a lot of the things we spoke about. Choudhury was giving specific examples about specific transactions the Blue Jays made and I remembered the transaction and the immediate reaction to the transaction, which helped keep the conversation going. Nakahara is not as involved in the decisions made in the front office but works with data that will help the Phillies on the field. My use of statistical models in my statistics classes gave me a general idea of the process Nakahara and his team takes to develop and read the models.

After the interviews, I asked both Choudhury and Nakahara if there was any way they could connect me with either a coach or player. Both told me to go through the media manager again, but I did not receive replies. The information I had felt like it was more than enough, but I wanted more. I knew there was really nothing else I could do to get in contact with more people, so I needed to give up on that end of things and begin the writing process.

For me, the first step in the writing process is crafting a headline. That is different for everyone, but I like to have a headline that I think will fit the story. As I write, the main theme of the story could change, so I would have to change my headline at the end. I like to hold off on that because sometimes I can get writing about something that gets away from the main point of the story. In that case, having a headline already in place will get me back on track and refocus me on the main idea of the story.

After the headline, the lead is the most important part of the story. If a story does not have a good lead, people will not read the story. There are two main types of leads: an inverted pyramid lead and a narrative lead. An inverted pyramid lead is typically used for hard news stories. Those leads give the most important information first in as few words as possible. You want to give the majority of the important details in the first few sentences. In a narrative lead, like I used in this story, you can be more creative and descriptive. You can set the scene for what you are about to write about. In this story, I tried to simplify the lead as much as possible. Baseball is a numbers game. In essence, that is the main idea of the entire story in four words. While it seems like that is too easy or a way out, it was easy to build off that idea.

After the lead is done, a story usually just flows for me. There are clear transitions to make and paths to go down. However, I usually do not have as much knowledge on a topic, information to work with or creative freedom as I had in this story. That is part of what made this

so fun to do. I did not have to leave out any details that I did not want to and could make it as long-form as I wanted. That made it really easy to write the first section of the story, as I could just go. Also, I found myself trying to cite sources for information that many people do not have off the top of their head. That seemed counterproductive and like a waste of time, but the sources I was using gave some extra information and are databases that readers could play around with to see different data.

The next part of the story incorporated the two like sources, Choudhury and Nakahara. Their experience helped me hammer home the point that analytics are an integral part of baseball. Through Choudhury's work with the general manager and Nakahara's work with the coaching staff, both of them gave very specific examples of how numbers impact the product seen on the field.

Lastly came the conclusion. To me, this is just as important as the lead. A good conclusion can resonate with someone and make them feel the impact of what they just read. The conclusion was one of the toughest parts of the story to write because I felt like there was so much more to write about. However, the place I ended was a good place because it would have turned into rambling without any more good quotes. Ending the story with the same idea you start it on can be a good way to ensure that the story comes full circle. If you can do that without making it sound awkward or abrupt, then you have successfully written a narrative story.

After the writing process is over, the editing process begins. I gave my story a read over before I sent it to my advisor to review also. There was very little that I caught on my end and my advisor had even less to go over. We redid the lead, which was where the idea of starting the story with, "Baseball is a numbers game," came about. It made perfect sense and made the story come completely full circle because the first and last sentence are now almost exactly the same.

Reflection

To say this project has tested my mental stamina would be an understatement. There were many times throughout the process where I felt like giving up, but the prospect of finishing with a product that I would be proud of kept me going forward. In the end, I was able to persevere and do what I needed to do, which is something I am very proud of.

The COVID-19 pandemic has taught me a lot about the journalistic process and about staying persistent. As I explained in the methods section, the pandemic threw a massive wrench in my plans. The entire world was put on pause for a while, which did not help my project. It took away a lot of the momentum I had gained from securing a day with the Syracuse Mets, but also put me into a tough spot. If the COVID-19 pandemic never happened, I feel like I would have been able to do things exactly how I had planned before and been able to execute my plan nearly flawlessly.

Since things did not go my way, I felt like I needed to devise a whole new plan. In reality, I just needed to be patient and trust the process. I waited until after the summer ended to contact Smorol again because I felt like that would be my best chance at getting a good response with a chance for him to help. During the summer, he had a lot going on with the season up in the air and people unsure of what the major league season would look like. However, once everything was underway, he likely had a lot off his plate.

Being told he could no longer help felt like a real gut punch. Not only did that leave me without a solid contact in the industry, but it also left me without a real sense of direction. I have been denied interviews before. It happens and it is something you deal with. As a writer for *The Stylus* for three years, I have been in the situation where a story has fallen through because I did not get any sources or important sources did not respond. The thing I have learned from this is

that you cannot panic and need to work with the information you can get your hands on. However, in this case, it felt like I would be unable to rebound because of how high I set my expectations.

After I was able to refocus and realized that I needed to push through and come up with another plan, I came across an Instagram page that gave me an idea of how to move forward. The page is run by a journalism student at Purdue University and is meant to promote the podcast he created. I followed the page because the podcasts were of interest to me but ended up connected with the creator and learned about how he went about securing his guests. He was the one who gave me the idea about contacting players over social media, so it was a new sense of hope and a new plan I had formulated.

The podcaster's success made me think that it would be just as easy for me to get in contact with people who are verified, but I was sorely mistaken. We were working with athletes from different sports, but I figured the easy-going nature of baseball players would play to my advantage. Again, I was wrong and disappointed. While I should have probably tried for players who were farther down the ladder in the minor league system, they do not have the same type of access to data and analytics like major leaguers do.

My final plan, the one that came to fruition, was the one that taught me the most. I am very familiar with the journalistic process. I know that sometimes things do not go your way and you need to just continue to power through the adversity to finish the story you were assigned. I assigned myself this story, so I needed to finish it no matter what. Sometimes, that requires you to take a shot in the dark. If you take a shot and bet on yourself, you have a better shot at success than if you sit around and feel bad for yourself.

After I sent the emails to all the major league teams, I checked my email probably 10 times a day. It only took a few days for an answer, but it felt like a year. Once I was set up with Choudhury and Nakahara, it was a tremendous relief off my back. I knew there was still a lot of work to be done, but I still felt like securing any interview was overcoming a major obstacle. I had hopes of being able to salvage my initial idea of talking to both executives and players but knew that when I was directed to Choudhury and Nakahara that it would be unlikely. Because of this, I did not have an exact idea of what the story would look like before I went into the interviews.

Like I mentioned in the methods section, the interview did not go the way I was used to. Despite that, I thought both interviews went well. My preparation consisted of some research on both Choudhury's and Nakahara's job, but I still wanted to ask them because I figured they could give me a more in-depth explanation as to what they do on a daily basis and how the help and impact the organization. I had three takeaways from the interviews that I can use as learning moments.

As the actual interviews were happening, I felt like I did a good job of keeping a good flow and not bouncing around between topics. When I listened back to the recording, I did not think I did as good of a job of that, but it was not bad. If I could have settled myself down and treated it like every other interview then I likely would have done much better with this.

Second, I definitely stuttered a lot when I was asking questions. When I go into interviews unscripted, I tend to have a more casual conversation like you are taught. I was so concerned about making sure I asked all the questions I wanted to, and I also did not want to ask them word-for-word how I wrote them when I was doing my research. When I went to ask the questions, I did not ask them as clean as I usually do and felt a little bit of pressure.

A final thing I should have done better was check the audio levels as the interview was going on and make sure they were always good. Because the interviews took place over Zoom and Choudhury and Nakahara both do not play, I determined that video would not help tell the story. It would just be a video of the interviews, which is not all that interesting and would take away from the writing. Also, I recorded the audio on my iPad. I lined up the speaker on the computer to the microphone on the iPad to ensure that the audio levels would be as good as I could get them. When I was interviewing Choudhury, I did not worry too much about the audio levels because I could hear him clearly the whole time. Nakahara was doing the interview with earbuds in and he was not as loud as Choudhury. I thought about this as a problem when the interview was taking place, but I did not want to stop to say anything. I should have, as a lot of the things he said were very quiet and difficult to decipher when I was listening back to the audio file. It did not turn out to be a big deal because I could pull good quotes that I could fully hear, but I need to pay closer attention to that from now on.

The actual writing process went very smooth for me. I was so invested in the story and had so much knowledge on the topic that it went very easily. I found myself citing sources so other people knew where to find the information, but I had a lot of the information already in my head. The strategy I like to use for writing a longer story like this is to give the background to the topic before getting into the meat of the story, like in a research paper. However, the writing is more concise in a story, so I feel like I was able to talk about the topic more than usual while still keeping a reader interested.

Like with the interviews, I had two takeaways that I can learn from with my writing process for this story. For one, I really took my time working with the transitions and made sure everything read well. It would seem like a given that a writer would want their story to read well,

but I think that the transitions are the most vital part of a story's readability. If ideas keep changing sporadically, then readers can get confused and stop reading the story. I thought I did a really good job of tinkering with the tougher transitions and really making sure everything had a nice flow to it.

My second takeaway is that I should not be concerned about a story's length. There was one point where I thought I had too much background information and needed to move on to introducing my live sources. I reasoned my way out of this by telling myself that the live sources would not be as impactful if all the background information was not there. It is not about the length of the story; it is about the story making sense and coming full circle.

Undertaking this project has taught me a lot about myself as a journalist. I learned a lot about changing my idea and working around obstacles. I did have to face a lot of difficulties and doubted my ability to finish the project a lot, but I am glad I pushed through and got it done by my deadline because I plan to be working on deadlines for the rest of my life. The fact that I was able to produce a product that I am truly proud of on top of that gives me a great sense of accomplishment.

Creative Work

A numbers game: How analytics are affecting decision-making in professional baseball

By Paul Cifonelli

Baseball is a numbers game. Baseball has always been a numbers game. As the game has progressed, so have the type and amount of numbers used. The one thing that has not changed is that players are sometimes evaluated strictly by the statistics they produce.

The best hitters in the early-1900s were those with high batting averages and RBI totals. Once the home run became popular in the mid-1900s, run prevention became essential and earned run average (ERA) and walks and hits per inning pitched (WHIP) came into play. The late-1900s had more all-or-nothing approaches from both hitters and pitchers. Players with the most home runs and strikeouts became the stars of the game, a trend that continues into today.

In 2003, author Michael Lewis published "Moneyball: The Art of Winning an Unfair Game." The book, which was turned into a movie in 2011, chronicled the 2002 Oakland A's and the team's unconventional process of acquiring players. Oakland was and is notoriously frugal and needed to replace its best players it could no longer afford. The team took an analytical approach to roster construction, using analytics and sabermetrics to find the best players for minimal value. The A's turned that into a division championship and postseason berth despite spending just over $40 million on its major league team, according to USA Today's player salary database. According to the same database, the New York Yankees' four highest paid players in 2002 made over $48 million. The Yankees also made the playoffs that year.

Since Oakland's success using "moneyball," nearly every other major league team has attempted to employ it in one way or another. Whether it is a team cutting costs and thinking for the future at the cost of being bad for a few years or getting multiple players who excel at one or two skills

instead of paying for a star, MLB teams have been taking ideas and concepts from the A's since 2002.

Two teams who have taken different approaches to team building in recent years are the Toronto Blue Jays and the Philadelphia Phillies. The Blue Jays made their first playoff appearance in four seasons in 2020 with a relatively young roster. Toronto has 14 players on its 40-man roster under the age of 25, while only five players on the team have had their 30th birthday. The Phillies, on the other hand, have been big spenders in free agency in recent years despite not making the playoffs since 2011. Philadelphia has 16 players under 25 years old and five 30 or older, similar to the Blue Jays.

The difference in the two clubs is how they spend money. Both have had minimal success in the last five years and have very similar roster construction, but the Phillies' 2020 payroll was seventh in the league at $80.893 million, according to Spotrac. Toronto spent $54.997 million, about $6.2 million less than the MLB average. Philadelphia's highest paid player is outfielder Bryce Harper, who signed the largest contract as a free agent in MLB history with a 13-year, $330 million deal. The Blue Jays' biggest earner is pitcher Hyun-Jin Ryu, who is there on a 4-year, $80 million contract.

This disparity in money spent between two teams in similar situations could come down to one of two possibilities. Either the team has a lot of financial flexibility or they are actively trying to keep payroll low by using "moneyball" techniques.

Two years ago, the same offseason Harper's contract was signed, Phillies owner John Middleton said the team had money to spend, according to NBC Sports Philadelphia.

"We're going into this expecting to spend money," Middleton told USA Today at the owners' meetings. "And maybe even be a little bit stupid about it. We just prefer not to be completely stupid."

While Toronto has not been going out and spending crazily like Philadelphia, the 2020 offseason may be different. However, to get here, the Blue Jays made more short-term signings for players who did not require much money.

Sanjay Choudhury is the Blue Jays' assistant director of research and development. He was hired by the organization at the beginning of the 2016 season, the last time Toronto made the playoffs before this year. Choudhury is helps with talent evaluation, whether it is for players to sign in free agency, players to select in the upcoming draft or players who will be facing the Blue Jays in the coming days. While the organizational shift from contending to rebuilding seems like it should have changed his job and how the organization looked for players, it did not.

"The organizational strategy changes and the way you approach things changes, but a lot of my job is still, 'Do you think player A is better than player B?," Choudhury said. "Whether you're a win now team or a win in three years team, you still have to get that decision right. We signed a player in 2017 we knew he wasn't going to impact our major league team that year. That's a decision that ended up having ramifications two years down the road, three years down the road and hopefully for our team for a few more years depending on if he's able to sustain that level of play he's had."

When looking for players to sign, management used to rely heavily on professional scouts' opinions from watching games. Some basic statistics that were attainable were used, but scouts were held in high regard. Now, with advanced metrics and analytics available at the click of a button, many people likely assume that scouting is dying. Choudhury says that is not the case.

"Maybe 20 years ago, before I was involved in the game, you had the 'nerds' who were just looking at performance," Choudhury said. "Where now you have 'moneyball,' where you're just looking at mostly stats. Scouts sometimes are very heavily invested in subjectivity and things that are more difficult to quantify. I actually think that those two roles, in terms of how you evaluate players, are closer together. Where someone in my position is expected to be able to evaluate players subjectively and also be able to think about how to quantify things like athleticism and makeup."

With scouts and analysts working so closely together now, both have a seat at the table with the decision-makers when it comes time to build the roster. Choudhury said that player transactions are not taken lightly and everyone who has valuable information on a player is allowed to present it before a decision is made.

"We have a group of people who give their opinions," Choudhury said. "I'm fortunate to be a part of that group. It's not a small group because [Blue Jays general manager] Ross [Atkins] is someone who's very collaborative and very open-minded and wants to get all the different opinions and viewpoints. It's going to be 15 people or more who are all going to weigh in, all going to see where they're at. At the end of the day, it's Ross' decision, he's the G.M. If there's 15 people who say we shouldn't do this, and he wants to do it, then he can do it. He usually doesn't operate that way. He's a very consensus-driven guy."

Alex Nakahara is the Philadelphia Phillies' senior quantitative analyst. His job is similar to the one Choudhury has with Toronto except he works a lot more on the day-to-day analytics that will be useful on gamedays. The data that is used goes through multiple departments before it is ready to be analyzed and broken down.

"Our process in a normal year is that the research and development staff create a lot of the models and produces things for the day-to-day," Nakahara said. "Then the advanced scouting analysts do the running of the models and look at the results and look at each individual opponent and see how they want to tailor it and customize it."

From there, analysts like Nakahara will look at the models and determine what the numbers mean and why they came out that way. In a year where COVID-19 is not limiting the number of people at a facility, Nakahara and the coaching staff would have conversations about the data from time to time to see what things they could change on the field or if something in the data seemed inconsistent with what was happening on the field.

Nakahara also mentioned that he does not have conversations with players about the data he breaks down. He has those conversations with the coaching staff and lets the coaches relay the data to the players.

When creating the models and breaking down the results, it yields things that can be done in certain scenarios. Seeing those results play out on the field can make the job rewarding, but Nakahara said there is also a downside to running so many tests.

"There definitely are fun aspects of it, especially the creative brainstorming part of it," Nakahara said. "There are some downsides, like we came up with 20 ideas but when can we do all of them? It's just not possible to do all these different things. Sometimes you can have too many ideas."

It has always been impossible to watch baseball without numbers being at the forefront. From batting average to ERA to home runs, players are judged by statistics. However, numbers and statistics are more involved in the game than ever before. From the executive decisions regarding

roster construction to coaches' decisions, analytics are involved in baseball more than anyone on the outside knows. Baseball is truly a numbers game.